Cornell University

Developing a Virtual Vault for Pseudopotentials: A NNIN/C Initiative

Derek Stewart

Cornell Nanoscale Facility

stewart@cnf.cornell.edu

Nanoinformatics 2010: 11/4/2010

CNF Cornell NanoScale Science and Technology Facility

NSF

Modeling across the NNIN

The National Nanotechnology Infrastructure Network (http://www.nnin.org)

NNIN Computation (NNIN/C)

Main Computation Nodes

Cornell (CNF)

Harvard

Stanford

Michigan

- **Computational mission**: *Develop modeling resources that complement and expand on the current experimental capabilities of nanoscale user facilities.*
- Provide access to on-site computing clusters.
- Provide a robust suite of simulation tools
- Construction of new codes to address user needs.
- Embedded simulation expertise to help develop strong research that combines experiments and simulations.

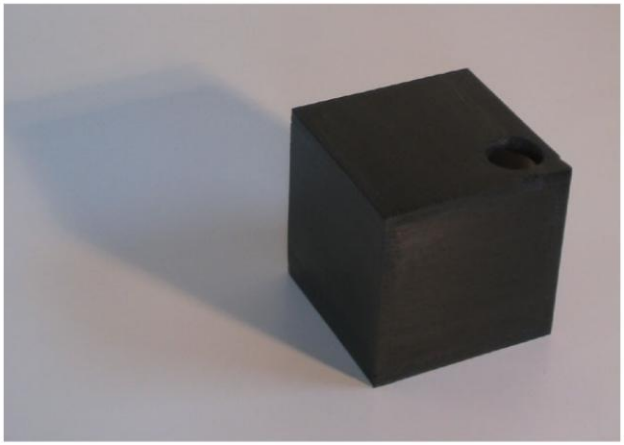
Building Trust in Nanoscale Simulations

The danger of the black box!

Garbage in / *Garbage out*

As a larger community begins to use simulation tools to model nanostructures, materials, and molecules, it is important to develop measures and standards to insure the quality of research.

Funding often encourages innovation not necessarily verification



Questions of scientific accuracy have already been raised in other simulation communities:

"Error… why scientific computing does not compute", Z. Merayi, *Nature*, **467**, 775 (2010).

"Accessible Reproducible Research", J. P. Mesirov, *Science*, **327**, 415 (2010)

"Reproducible Research", Yale Law School Roundtable, *Computing in Science And Engineering, **Sept/Oct Issue** pg. 8-10, 2010*

Developing a Virtual Vault for Nanoscience

A NNIN/C cyber-infrastructure resource for the nanoscale community

The Virtual Vault will provide two main resources:

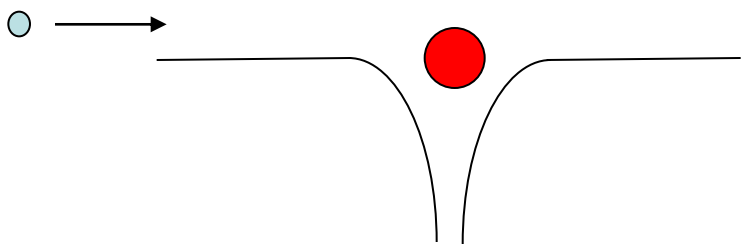
A **Pseudopotential Database** which will provide thoroughly vetted pseudopotentials along with associated data on key calculated material properties.

A **Input File Bank** that allows researchers to upload input files and associated data from NSF funded research and publications

Pseudopotentials

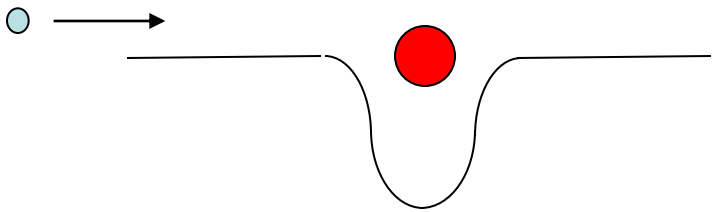
A key component for ab-initio approaches

Real Potential



$$V_{ION} = -\frac{Ze}{r} \qquad r \to 0, V_{ION} = -\infty$$

Pseudopotential

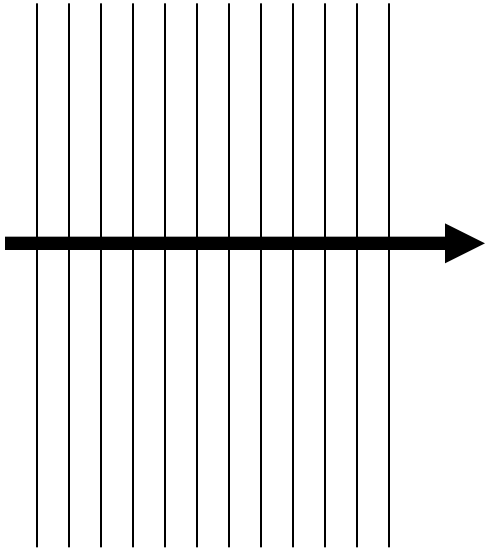


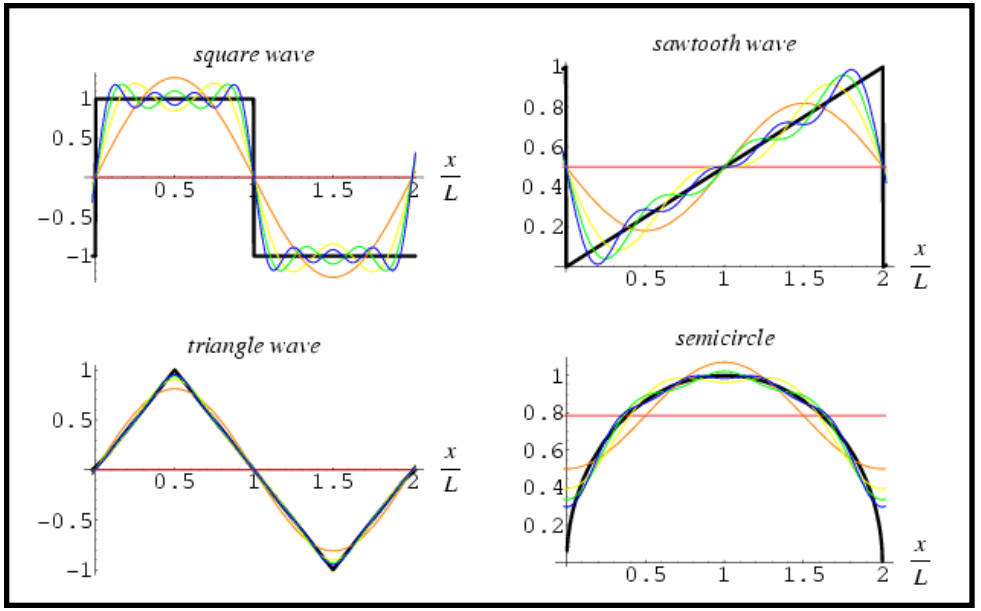
Pseudopotentials can accurately describe electron scattering from ion cores and allow us to predict material properties using a much shallower potential around the ion.

This leads to **faster calculations** and the chance to look at **bigger systems.**

Used in VASP, ABINIT, CPMD, Quantum Espresso, PARSEC, and many more codes.

Plane Waves



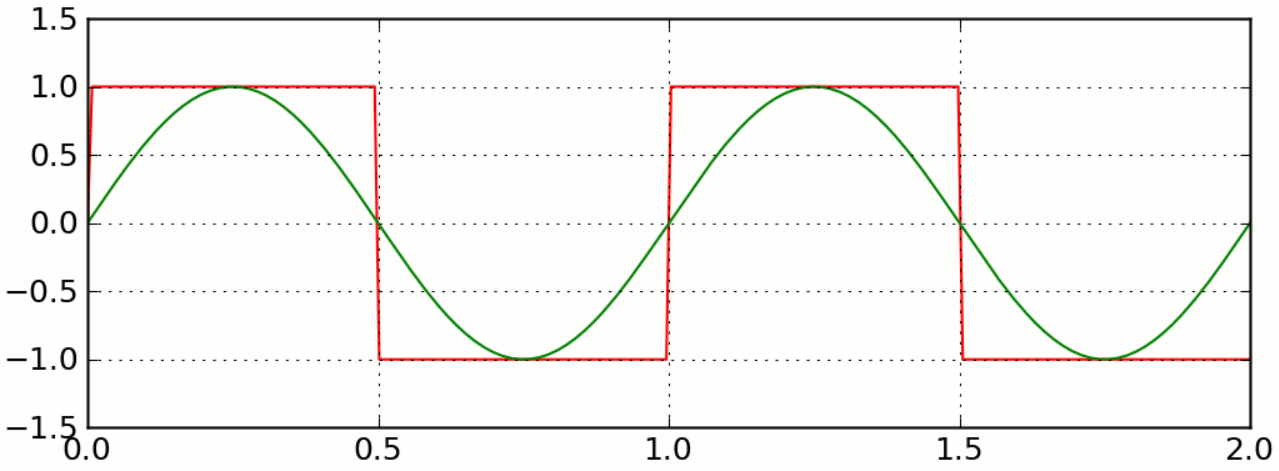


Numerous codes use a plane wave basis set?

WolframMathWorld

1. Systematic improvement in the quality of the calculation (*just add more plane waves*)
2. Makes evaluations of gradients and lapacians very easy
3. Basis set is independent of the atomic positions (no Pulay terms in the forces)

Describing a Square Wave with Fourier Components



It is difficult for plane waves (sin and cos waves) to describe sharp features like a square wave.

Similarly for electronic structure, some pseudopotentials that are fairly smooth and shallow can be described with fewer plane wave components, leading to faster calculations.

http://en.wikipedia.org/wiki/File:SquareWave.gif

The Current Problem: Which pseudopotentials do you trust?

- *Rolling your* own - Development of a robust pseudopotential is a time-consuming and specialized process
- Online databases of pseudopotentials are available but they contain little or no supporting data to verify their accuracy and they often have gaps (missing potentials for certain atoms).

Researchers are forced to verify the same atomic pseudopotentials through a series of calculations on simple materials or else they neglect this step **at their peril** !!!

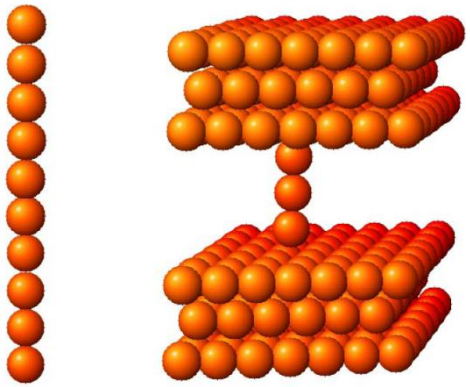
Case Study: Magnetic Forms of Palladium

Recent evidence for magnetic forms of palladium.

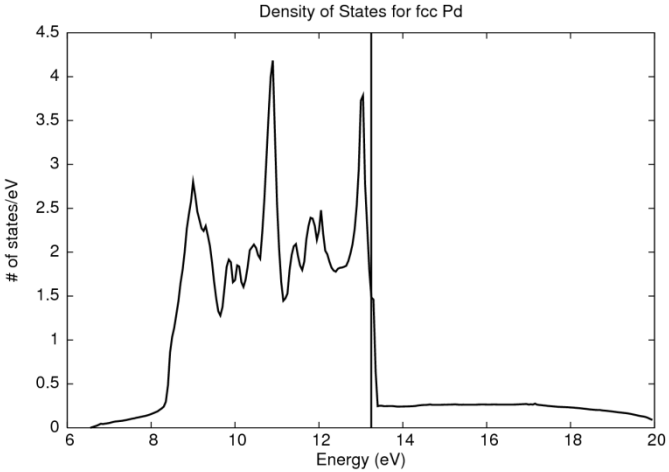
•**hcp Pd films** (~100 monolayers) show magnetic moments of 0.16 μ_B/atom (Hüger & Osuch, Europhys Lett. 2003)

•**Pd clusters**: weak ferromagnetism observed in Pd clusters 2.5 nm in size (Sampedro *et. al.* PRL 2003)

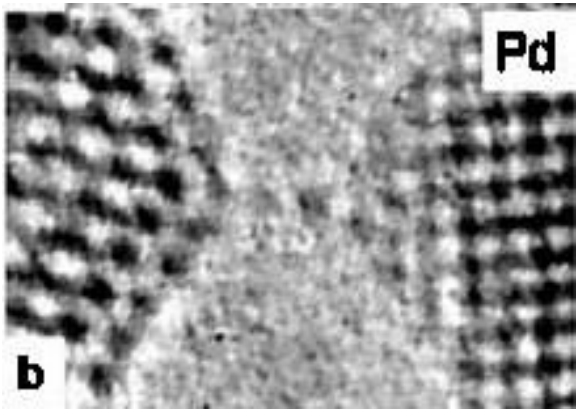
•**Pd break junctions and nanowires** – experimental work and theoretical calculations indicate that Pd atomic strands could be magnetic.



Magnetic moment 0.7μB/atom in Pd atomic strands, A. Delin *et al.*, PRL, 2004

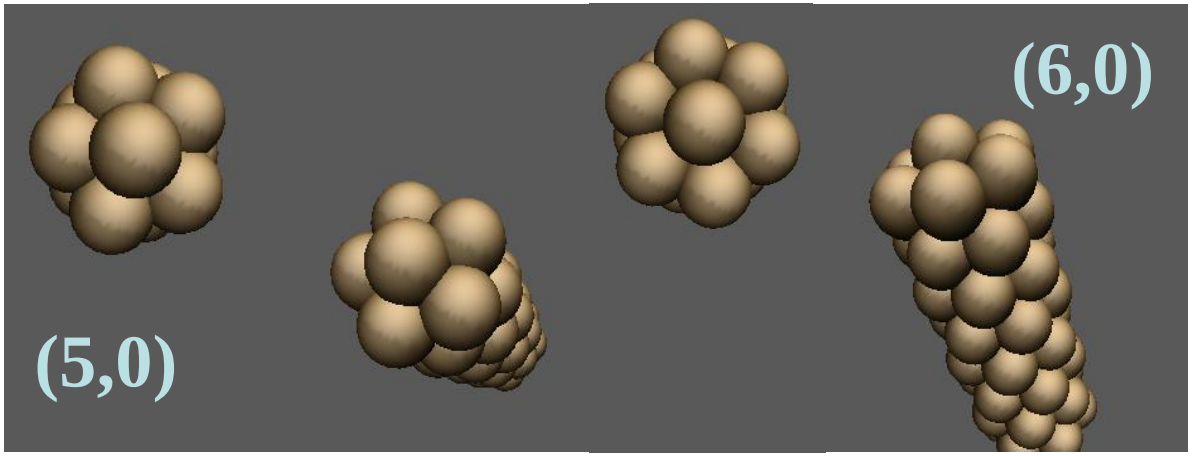


Bulk fcc Pd sits at the non-magnetic edge of the Stoner criterion



Magnetic transport through nanowires in Pd break junctions Rodrigues *et al.,* PRL 2003

Are Coaxial Palladium Nanowires Magnetic?



In 1998, Gulseren, Ercolessi, and Tosatti predict that ultrathin metal wires will have non-crystalline structures based on computer simulations.
(PRL, **80**, 3775 1998)

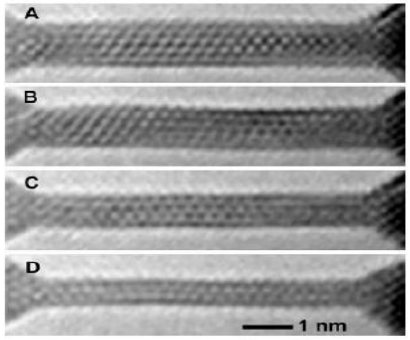
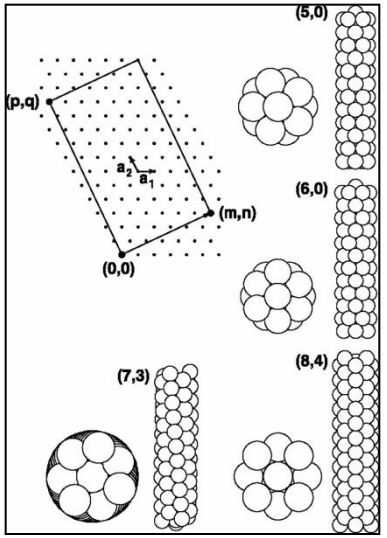


Fig. 1. TEM images of stable gold nanowires observed during one thinning process. The diameters of the wires in (A), (B), (C), and (D) are 1.3, 1.1, 0.8, and 0.6 nm, respectively. The dark dots represent positions of atoms projected on the image plane. The dark dots are aligned on atom rows along the wire axis. These wire images are wavy, particularly in (D).

Gold coaxial nanowires (Kondo & Takayanagi, Science, 2000).



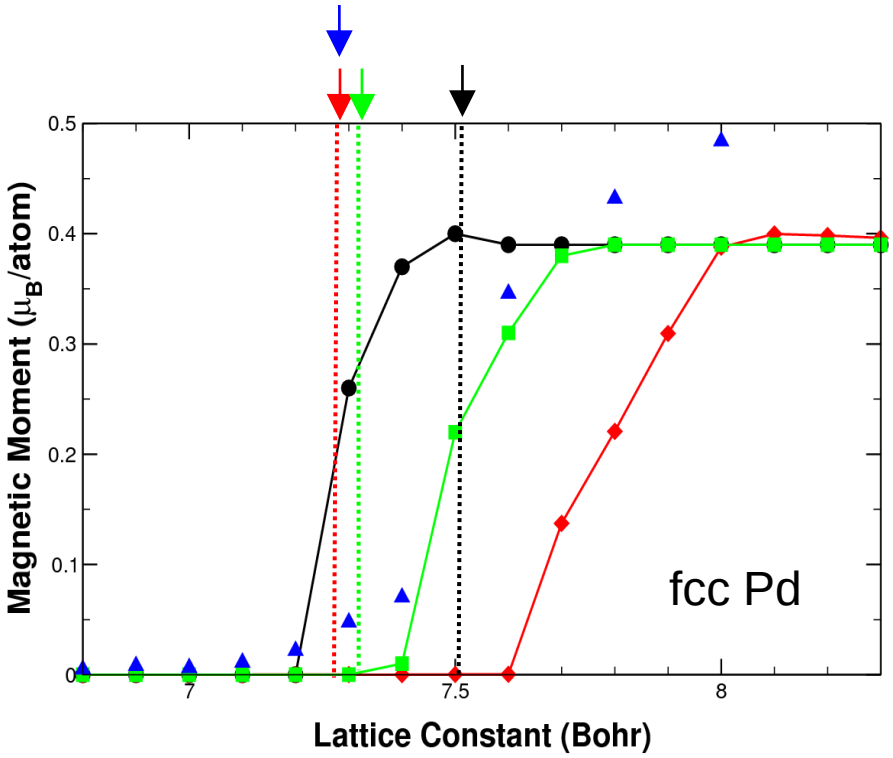
Coaxial Nanowires are formed by wrapping a (111) sheet

Problem! Most pseudopotentials predict that all forms of Pd are magnetic!

Total energy calculations with structural relaxations were done using the plane wave density function package Quantum Espresso (formerly known as PWscf), Abinit, and LMTO.

Several pseudopotentials were examined for this study:

LDA pseudopotential with non-linear core correction and a semi-core d state was the only one capable of reproducing the correct non-magnetic fcc Pd bulk state.



LMTO (LDA – all electron)
PWSCF (LDA PSP-semicore d)
PWSCF (GGA PSP -semicore d)
ABINIT (LDA PSP)

A CyberInfrastructure Solution

The NNIN/C is developing a database for common ab-initio codes that will provide thoroughly vetted pseudopotentials.

Each pseudopotential will have additional data on common calculated material properties (lattice constant, bulk modulus, phonon dispersion, etc) and references to publications that used these pseudopotentials

Researchers can also submit their own pseudopotentials. These pseudopotentials will be automatically tested on calculations for common materials and the results will be published online with the pseudopotential.

This process will help save time and improve research quality.

Phase 1: Clearinghouse

Currently there is *no* central online resource for pseudopotentials online.

NNIN/C has develop a clearinghouse for pseudopotential data:

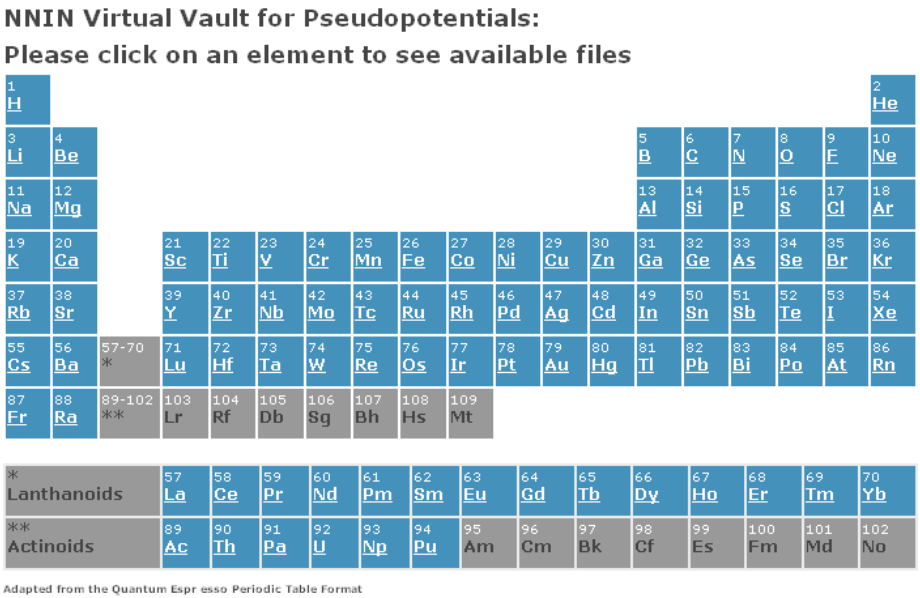
- Access to 12 online pseudopotential collections (Abinit, Quantum Espresso, Vanderbilt, CASINO, etc)
- Access to 10 pseudopotential generators
- Access to 4 pseudopotential convertors
- Key Pseudopotential References for New Users

Phase 2: Online Database

NNIN/C Pseudopotential Database

- The **first** centralized database for pseudopotentials from multiple codes
- Currently **~800** pseudopotential files available (ultrasoft, norm-conserving, LDA, GGA, etc) from Abinit, Quantum Espresso, Qbox, Bennett-Rappe)
- Web-accessible and searchable SQL database with a PHP web interface to dynamically generate pages.

Pseudopotential resources in the Virtual Vault for Nanoscience



NNIN/C Pseudopotential website provides a clearinghouse for online resources:

- Links to **11** online pseudopotential databases (Abinit, Quantum Espresso, Vanderbilt, CASINO, etc).
- **12** Pseudopotential generators
- **4** Pseudopotential convertors
- Key Pseudopotential References

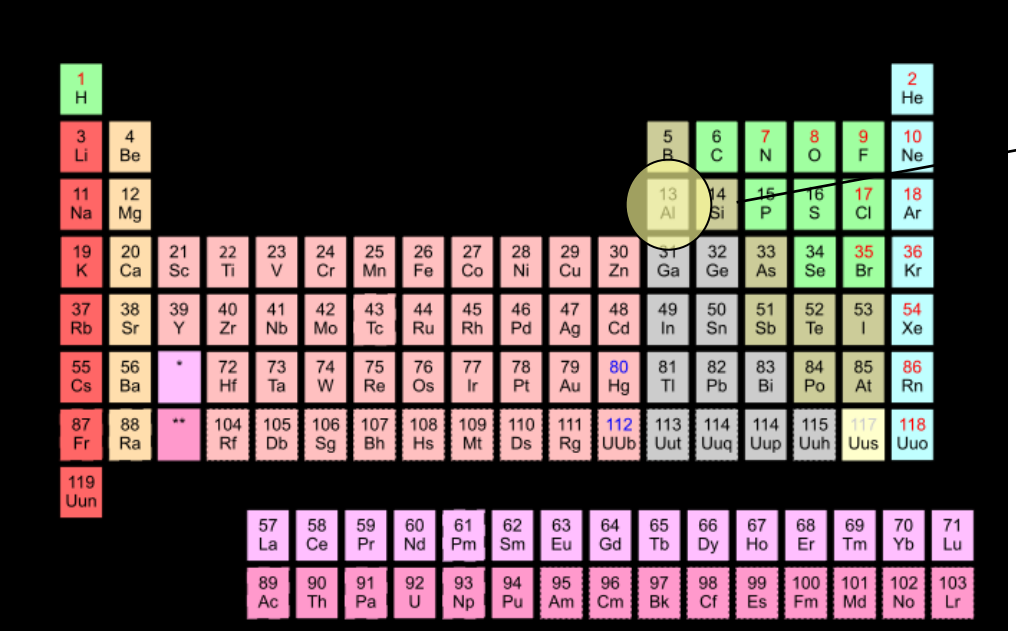
NNIN/C Pseudopotential Database

- The **first** centralized database for pseudopotentials from multiple codes
- Currently **800** pseudopotential files available (ultrasoft, norm-conserving, LDA, GGA, etc)
- Web-accessible and searchable SQL database with a PHP web interface.

Phase 3: Interactive Database

The end goal is to provide a truly interactive platform for pseudopotentials:

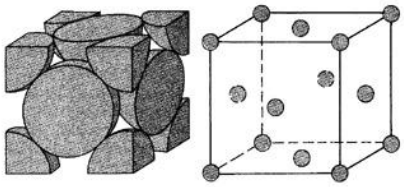
- A vast collection of *verified* pseudopotentials with relevant predicted material properties
- The ability for users to *upload* their own pseudopotentials for evaluation
- Easy *translation* between pseudopotential formats
- Links to publications as well as comment spaces



Pseudopotential: **13al.981214.fhi**

Al Z=13

fcc crystal structure



Calculation Details:

Code: ABINIT

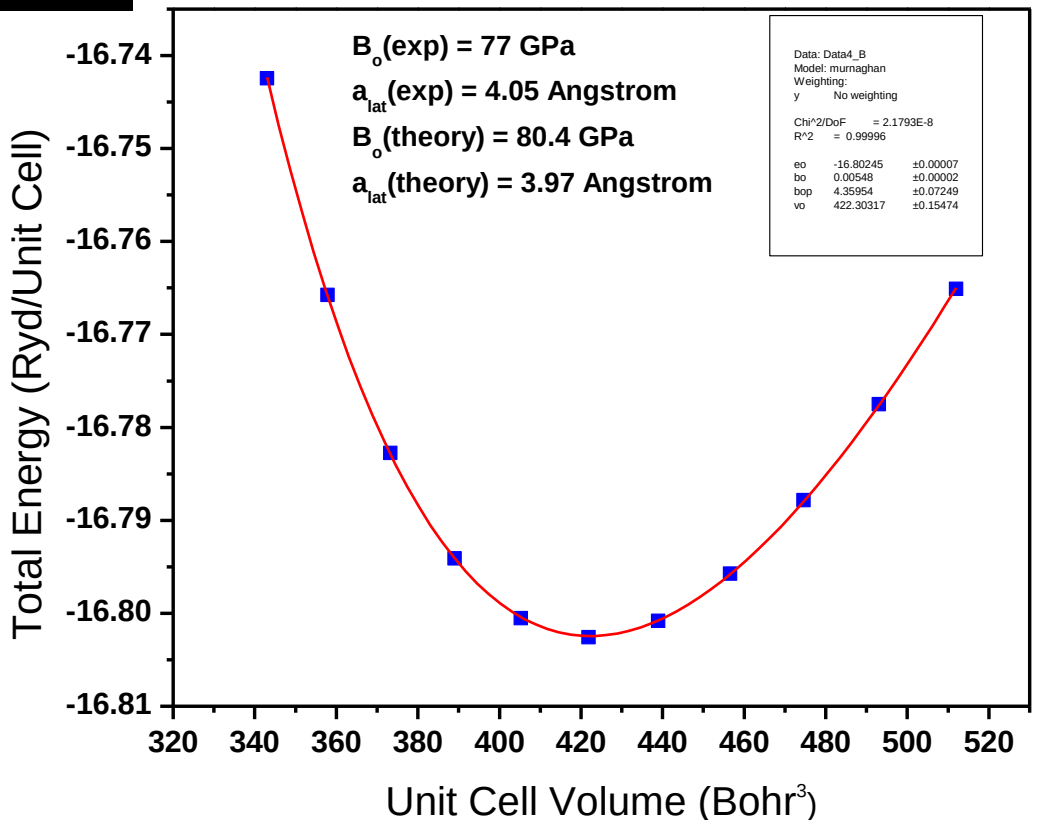
Energy Cutoff: 60 Ha

K-Point grid: 16x16x16

Smearing Factor: 0.05 Ha

Publications that used the potential:

M. Giantomassi *et al.*, Phys. Rev B. **72**, 224512 (2005)



Input File Bank

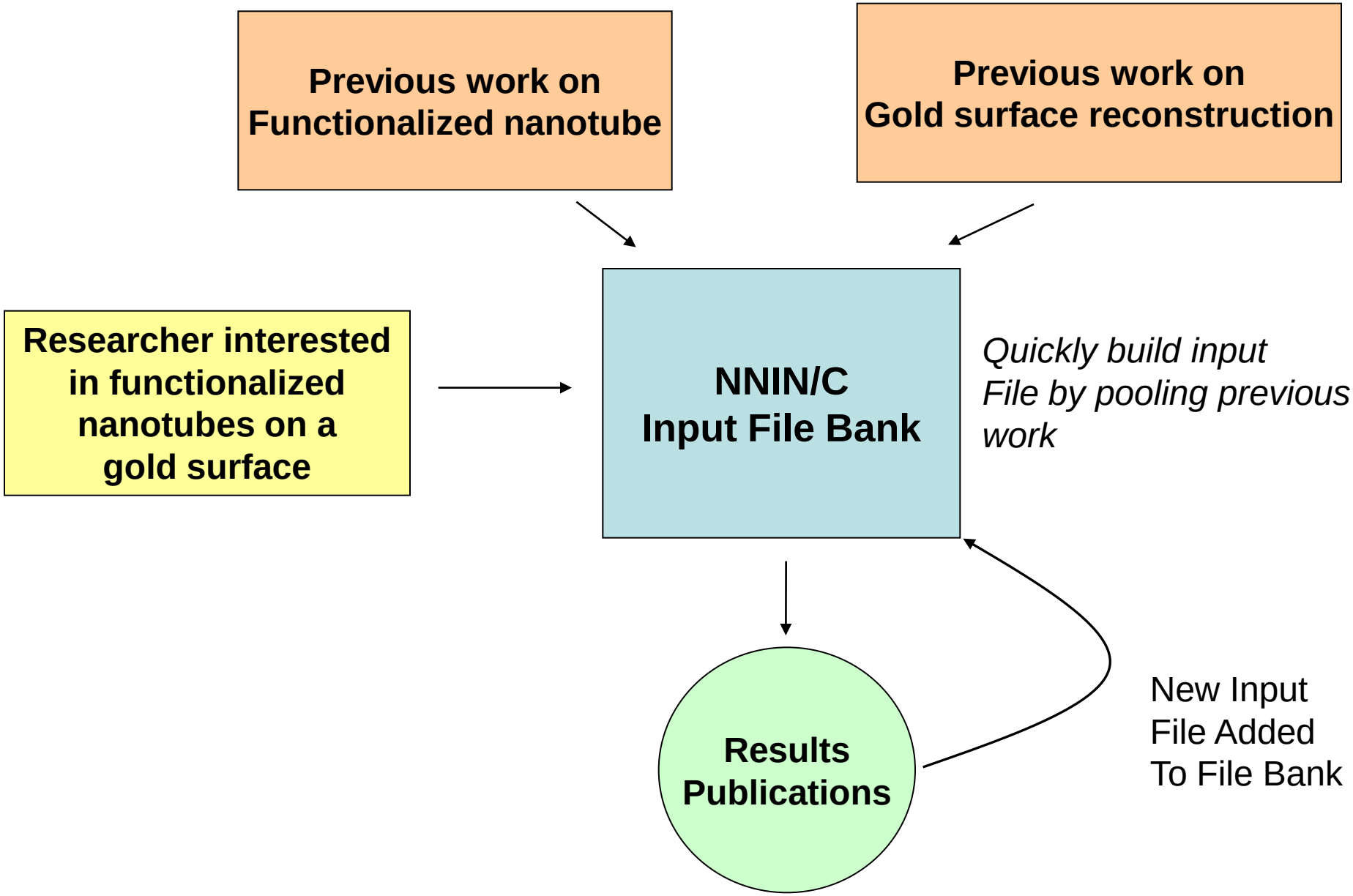
leveraging past research

The Virtual Vault will provide a location for researchers to upload input files and associated data from their research and publications.

This information will **help researchers collaborate** and also build on the success of previous efforts.

The NNIN/C will work with NSF, NIH, and major journals to encourage researchers to submit common input information to the Virtual Vault when research is published.

This will also help **improve scientific accuracy**. Making the input files used in published research can provide an important route for other researchers to repeat calculations and verify results.


Previous work on
Functionalized nanotube
Previous work on
Gold surface reconstruction
Researcher interested
in functionalized
nanotubes on a
gold surface
NNIN/C
Input File Bank
Quickly build input
File by pooling previous
work
Results
Publications
New Input
File Added
To File Bank

Community Challenges

How do you encourage researchers to share their data or link it with their publications?

Concerns that other researchers could leap-frog over their current work or simply write their paper based on the researcher's included data.

Fear that other researchers could challenge their papers based on their input files?

Possible solution: Only accessible to reviewers or else an embargo on release of input file data

Conclusions

A clear need exists for central resources that provide trusted components to help insure accuracy in nanoscale simulations and also encourage collaboration.

Two efforts of the NNIN/C:

Virtual Vault for Pseudopotentials

Input File Bank

These efforts grow with collaboration! If you are interested In helping out, please let me know.

Acknowledgements

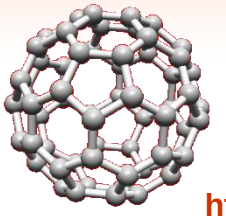
Richard Hennig (Cornell) – his group is working to translate CASINO pseudopotentials to Quantum Espresso format for inclusion in the database. They will also evaluate key characteristics of the pseudopotentials.

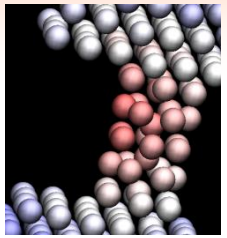
Derek Warner (Cornell) – for advice on online databases, PHP, and SQL.

Funding for the Cornell Nanoscale Facility is made possible through the NSF and NYSTAR.

2010 Joint NNIN/C & NCN Fall Workshop

Building a Collaborative Framework for Nanoscale Simulations





November 14th-16th, 2010
Cornell University, Ithaca, NY

http://www.cnf.cornell.edu/cnf_fallworkshop2010.html

Overview: Nanoscience relies on intensive atomistic simulations that generate copious amounts of data. Since various ab-initio and empirical approaches exist in the field, there is a growing need to develop a ***common collaborative framework*** for meaningful comparisons between calculations and to also provide a set of robust technical components and results for the next generation of researchers.

This workshop brings together leading figures in fields of materials and nanostructure simulation to help develop a set of *lingua franca* formats & libraries for easy translation of input and output files between codes. We will also address the creation of international cyberinfrastructure resources that provide trusted components for atomistic calculations.

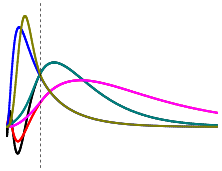
Confirmed Speakers:

Gerhard Klimeck (Purdue)
Geoffrey Hutchison (Pittsburgh)
Matthieu Verstraete (University of Liege, Belgium)
Karsten Jacobsen (Technical University of Denmark)
Francois Gygi (UC Davis)
Howard Sheng (George Mason)
Adri van Duin (Penn State)
Jim Sethna (Cornell)
Derek Warner (Cornell)
Derek Stewart (Cornell)
Xavier Andrade (Harvard)
Julian Velev (U of Puerto Rico)

Registration Deadline:
November 3, 2010
Seating is limited to 40 participants

Organizer: Derek Stewart (stewart@cnf.cornell.edu)

Workshop Topics



Trusted Components for Calculations

- *Pseudopotentials*
- *Empirical Potentials*
- *Basis Sets*
- *Crystallographic Data*

Standing on the Shoulders of Giants

- *Developing a Database of Input Files based on Previous Research*
- *Common IO Formats & Libraries for Ab-initio Calculations*

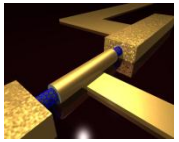
Overview of Current Efforts in the Field Including…

- *Atomic Simulation Environment*
- *OpenBabel*
- *NCN Nanohub*
- *NNIN Virtual Vault*
- *European Theoretical Spectroscopy Facility (ETSF) DFT Libraries*

Fully Reproducible Science

- *Linking Calculations directly with Publications*
- *Input Files as a part of Article Supplementary Info*

Tutorial Sessions on ASE, Abinit, Qbox, Avogadro, and more!



Photograph by Zack Schildhorn

NNIN Nanoscale Science, Engineering & Technology

NSF

NETWORK FOR COMPUTATIONAL NANOTECHNOLOGY